UDC 330.34. 351.862

PRACTICAL APPLICATION OF INTELLECTUAL DATA ANALYSIS IN THE CONTEXT OF NEURAL NETWORKS: A COGNITIVE-FREE APPROACH IN THE PROPAGANDA SYSTEM

VIVCHAR Oksana[1], SOBKO Olha[2], MUZHYLIVSKYI Nazarii[3]
[1] West Ukrainian National University
https://orcid.org/0000-0001-9246-2226
[2] West Ukrainian National University
https://orcid.org/0000-0001-8317-0563
[3] West Ukrainian National University
https://orcid.org/0009-0003-1974-4253

This article investigates the use of artificial neural networks as an intelligent data analysis tool. Due to the rapid growth of data volume caused by the automation of various technical processes, the use of such networks is particularly relevant. The article highlights the main methods of data analysis in the context of security science and analyzes the advantages and disadvantages of these methods. It is proven that in the system of practical application, the use of artificial neural networks for intelligent data analysis is quite appropriate. The areas of application of intelligent data analysis and existing systems are analyzed. A comparative characteristic of the use of neural network architecture in the context of security science conditions is carried out. An approach to a neural network for data selection with the separation of the corresponding stages is proposed. On the basis of which a practical representation of the simplest recurrent neural network with the separation of the corresponding stages is proposed. An algorithm for building the specified system is proposed. Promising directions for the use of intelligent data analysis methods are identified.

Keywords: data mining, neural networks, artificial neural networks, recurrent neural network, enterprise security, economic and mathematical model, propaganda.

https://doi.org/10.31891/mdes/2025-18-5

Стаття надійшла до редакції / Received 22.09.2025
Прийнята до друку / Accepted 16.11.2025

PROBLEM STATEMENT IN GENERAL FORM AND ITS CONNECTION WITH IMPORTANT SCIENTIFIC OR PRACTICAL TASKS

One of the effective tools for data mining has been mathematical statistics. Mathematical statistics methods have proven useful mainly for testing pre-formulated hypotheses. Therefore, there was a need to develop new modern methodologies for data processing and analysis. Data mining (DM) became such a new methodology. We draw attention to the fact that modern statistical methods are no longer able to constructively process large data sets. That is why data mining makes it possible to identify hidden connections in large data sets. It has become clear that the specific requirements for data mining processing include: handling data of unlimited volume; accommodating heterogeneous data types (quantitative, qualitative, and textual); ensuring the results are specific and comprehensible; and providing user-friendly tools for processing raw data.Thus, the concept of templates that reflect fragments of multi-faceted relationships in the data is the basis of data mining (discovery-driven data mining, Data Mining, DM). These patterns determine the patterns inherent in data subsamples that can be compactly expressed in a form understandable to a person. Pattern discovery is conducted using methods that are not constrained by a priori assumptions regarding the sample structure or the distribution types of the analyzed indicators' values. A key principle of IDA (Intelligent Data Analysis) is the non-triviality of the patterns being searched for. This means that the identified patterns should reveal non-obvious, unexpected regularities within the data, which constitute what is known as hidden knowledge.

The main reasons for the increasing relevance of IDA are: the rapid accumulation of data; the widespread computerization of business processes; the integration of the Internet into all areas of activity; advancements in information technology, such as improved data storage solutions; progress in various technologies, including the rapid growth in computer system productivity and storage capacities; and the implementation of Grid systems.

The algorithms used in IDA require a large number of calculations. Previously, this was a limiting factor for the widespread practical application of IDA, but today's increase in the productivity of modern processes has removed the acuteness of this problem. Now, in a reasonable time, it is possible to conduct a qualitative analysis of hundreds of thousands and millions of records. It has been established that IDA is an interdisciplinary field that arose and developed on the basis of such sciences as applied statistics, pattern recognition, artificial intelligence, database theory, etc. [3].

At the beginning of its development, the use of neural networks in data analysis caused mixed reviews due to such shortcomings as the complexity of the structure, too long a training period and poor

interpretability. But they were compensated by a set of positive qualities, such as a low error rate, constant improvement and optimization of various algorithms for training networks, the algorithm for obtaining rules, and the algorithm for simplifying networks, which make neural networks an extremely promising direction in the field of data analysis [2, p. 402].

The admissibility of a high level of data noise and a low probability of error, as well as the continuity of the network training process, constant improvement of its structure – all this gives grounds to consider the use of neural networks as one of the most promising methods of data analysis. After all, the use of neural networks allows you to analyze the state and identify the parameters of objects even in the case of incomplete data.

According to the conducted studies, it was established that an artificial neural network is a mathematical model of the human brain. The brain consists of a huge number of particles – neurons, interconnected by communication lines that transmit signals. Passing through neurons, signals change. The presence of an innumerable number of such neurons allows a person to solve complex problems, remember large amounts of information and make decisions [4].

The development of such a model enabled the computer to perform specific independent tasks. To train a neural network, a training sample is provided, allowing the network to autonomously adjust its structure for optimal modeling of the sample. After the training process, the network demonstrates the ability to produce calculation results using input data that was not included in the training sample.

We would like to highlight that data mining has a practical application in applied chemistry. IDA methods are widely used in both organic and inorganic chemistry. A common challenge in this field is understanding the characteristics of the chemical structure of specific compounds and their defining properties. This task becomes particularly crucial when analyzing complex chemical compounds, whose descriptions include hundreds or even thousands of structural elements and their interconnections.

Many more examples can be given where IDA methods play a major role. The peculiarity of this lies in their complex system organization. They belong mainly to the supracybernetic level of system organization, the regularities of which cannot be accurately described in the language of statistical or other analytical mathematical models. Data in the indicated areas are heterogeneous, heterogeneous, non-stationary and often have high dimensionality [1].

Thus, based on the material considered above, it can be argued that the potential of IDA gives impetus to expanding the boundaries of application of this technology in the modern world of computer technologies. Regarding the prospects of IDA, the following directions of development are possible:

– allocation of types of subject areas with corresponding heuristics, the formalization of which will facilitate the solution of the corresponding IDA problems belonging to these areas;

– creation of formal languages and logical means with the help of which reasoning will be formalized and the automation of which will become a tool for solving IDA problems in specific subject areas;

– creation of IDA methods capable not only of extracting regularities from data, but also of forming some theories based on empirical data;

– overcoming the significant lag between the capabilities of IDA tools and theoretical achievements in this area (Narkevich et al, 2021).

The main feature of IDA is the combination of a wide range of mathematical tools (from classical statistical analysis to new cybernetic methods) and the latest achievements in the field of information technology. IDA technology harmoniously combines clearly formalized methods and methods of informal analysis, that is, quantitative and qualitative data analysis [7, p. 90].

Most of the analytical methods used in IDA are well-known mathematical algorithms and methods. What is new in their application is the possibility of their use when solving certain specific problems, due to the new technical and software tools that have appeared.

An important advantage of using neural networks for processing data sets is a significant increase in the speed of the process compared to traditional mathematical methods, the ability to train a neural network using reference samples, as well as a change in the network typology (selection of input parameters that guarantee obtaining a model of the highest accuracy), based on the requirements of the problem being solved.

It has been established that artificial neural networks are mathematical models, as well as their software or hardware implementations, built on the principle of organization and functioning of biological neural networks – networks of nerve cells of a living organism. They arose in connection with the study of processes occurring in the human brain, as well as with the ability to predict them. An artificial neural network consists of many simple computational elements (neurons), interconnected in a certain way. The

most common are multilayer networks in which neurons are combined into layers, in turn, is a collection of neurons that receive information in parallel from other neurons in the network at each time step. In other words, the outputs of neurons are connected to the inputs of other neurons, allowing the signal from one element to be transmitted to another. Thus, neural networks are advisable to use for solving difficultly formalized tasks (which require labor-intensive calculations). Such tasks include classification of objects of economic analysis. It is proposed to solve the problem of data mining in three stages. In the first stage, it is necessary to determine the type of artificial neural network, in the second stage – the parameters of the network architecture (number of neural layers, etc.), in the third stage, the network is trained, that is, the network is given the necessary parameters that determine the correct output result, or adjusted at the initial stages, if the training takes place «with a teacher» [8].

Based on the conducted research, the structure of the neural network is proposed. Let us present the practical aspects of the neural network architecture. The type of network we have chosen is a recurrent neural network (RNN) for analyzing data sequences or a deep neural network (DNN) for classifying threats. Fig. 1. shows the basic structure of the layers of the neural network.

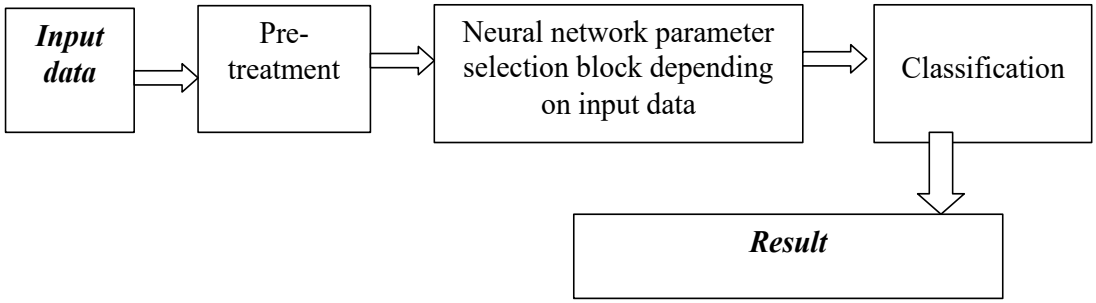

Fig. 1. Proposed Neural Network Approach To Data Selection

*created by the authors

Let us describe the structure of the neural network layers according to Fig. 1.

1. Input layer:

– input parameters: event logs (log files), network packets, system calls, threat signatures.

– the dimension of the input data depends on the type of task (for example, 10-20 key features after data preprocessing).

Data preprocessing. The input layer of an artificial neural network is responsible for receiving data for further processing. In the context of security, the input layer receives data from various sources, such as event logs (log files), network packets, system calls, and other structured or unstructured data that reflect the state of the system. This data is preprocessed to normalize, remove noise, and extract key features, such as event time, request type, IP addresses, packet size, or potential threat signature. The result is a set of fixed-dimensional vectors that represent relevant information for analysis. A feature of the input layer is its ability to adapt to different data formats through the prior use of mining techniques such as feature extraction or embedding. For example, text data can be transformed into a numerical representation using tokenization and vectorization, while network data is analyzed using filtering and clustering. In some cases, specialized approaches are used, such as one-hot encoding for categorical variables or time-series transformation for event sequences. This ensures high quality and relevance of the input data, which is critical for the accurate operation of the subsequent layers of the neural network [6, p. 63].

2. Hidden layers:

– Dropout layer to reduce overfitting;

– Conv1D/Conv2D layers (for CNN) or LSTM/GRU (for RNN) to process sequences and extract patterns;

– Batch Normalization to stabilize learning.

The hidden layer of a neural network is responsible for processing and analyzing input data in order to extract hidden patterns and create a high-level representation. In the context of security, hidden layers can contain different types of computational units, such as convolutional layers (Conv1D/Conv2D) to detect patterns in data, recurrent layers (LSTM/GRU) to analyze sequences of events, and fully connected layers (Dense) that combine information from previous layers. For example, convolutional layers can find specific features of anomalous activity in network packets, while recurrent layers analyze time dependencies associated with a series of system calls or events. Optimization mechanisms such as Dropout,

Batch Normalization, and activation functions (e.g., ReLU, Sigmoid) play a special role in the hidden layer. Dropout reduces the risk of overfitting by randomly disabling neurons during training, which allows the model to better generalize to data. Batch Normalization stabilizes the learning process by normalizing the input data of each layer, which accelerates convergence. Activation functions add nonlinearity, which allows the network to solve complex tasks such as classification or prediction. In the case of security tasks, hidden layers work on the integration and transformation of complex multidimensional features, which helps to identify even weak signals of potential threats. This makes hidden layers a central component in achieving high model accuracy.

3. Output layer:

– Softmax for classification (e.g., «anomaly» or «normal»);

– Sigmoid or ReLU for risk assessment (0–1).

The output layer of a neural network is responsible for generating the final analysis result. In security tasks, this layer transforms the processed data from the hidden layers into a user-friendly format, such as a threat classification (anomaly or normal state) or a risk assessment (in percentage). The output layer consists of one or more neurons, depending on the type of task. In the case of classification, for example, to detect the type of attack, the Softmax activation function is used, which converts the outputs of the neurons into probabilities for each class. For regression tasks, such as predicting the level of risk, the ReLU or Sigmoid activation function is usually used. A feature of the output layer is its adaptation to the specifics of the tasks. In the context of cybersecurity, it can include binary classification (for example, «attack» or «normal state») or multi-class classification to identify specific types of attacks [7]. In addition, the output layer can provide multidimensional results, for example, indicating both the type of threat and its intensity or potential impact. This makes the output layer not only a key component for completing the analytical process, but also an important tool for ensuring the usability of the model in practice, allowing it to be integrated into SIEM systems, monitoring dashboards or automated response tools.

We detail the practical stages of forming a convolutional neural network according to the specified main components (Fig. 2).

It should be noted that in the process of training on the identification of neural networks, the following components should be justified:

– algorithms used: Adam, RMSprop;

– loss function: cross-entropy for classification, MSE for regression;

– data for training: simulated attacks, real threat data.

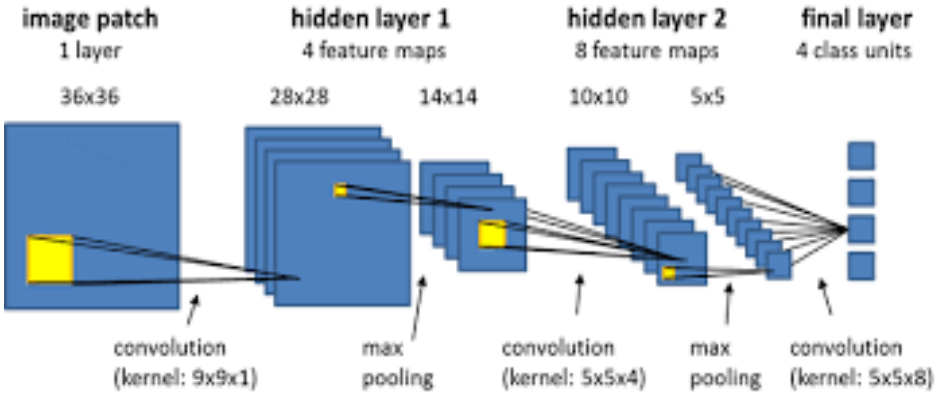

Fig. 2. Architecture Of A Convolutional Neural Network

*created by the authors

It is impossible to ignore the fact that the next component is the optimization of the neural network component, in particular:

– L2 regularization to avoid overtraining;

– model ensemble (using multiple networks to reduce the error);

– transfer learning approach to adapt previously trained models [5, p. 110].

The procedural algorithm for the operation of the simplest neural network:

1. Data preprocessing: normalization, noise removal, feature extraction.

2. Analysis stages: anomaly detection (example: unusual database queries); attack type classification (SQL injection, DDoS, Brute Force); risk prediction based on previous data (dis. Fig. 3).

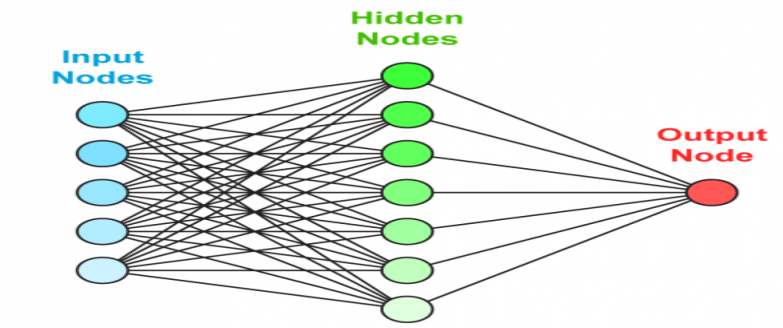

Fig. 3. Architecture Of The Simplest Neural Network

*created by the authors

The next and final component of the formation of a recurrent neural network is the results and discussion, namely:

– model efficiency: accuracy, recall, F1-measure, ROC-AUC;

– comparison with traditional approaches without neural networks;

– examples of successful threat detection on test data [4].

The visualization of the simplest recurrent neural network is shown in Fig. 4.

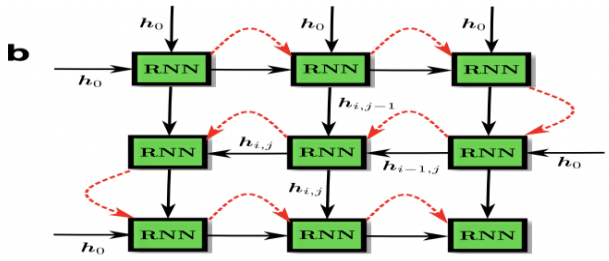

Fig. 4. Practical Representation Of The Simplest Recurrent Neural Network

*created by the authors

Therefore, it should be noted that the main advantage is the use of neural networks in the context of security science. It is recommended to apply practical application of network systems data, in particular implementation in SIEM systems. Promising directions for further research directions are the use of hybrid architectures.

CONCLUSIONS FROM THE RESEARCH AND PROSPECTS FOR FURTHER INVESTIGATIONS IN THIS AREA

Data mining systems are used as a mass product for business applications and as tools for conducting unique research (genetics, chemistry, medicine, etc.). IDA leaders associate the future of these systems with their use as intelligent applications built into corporate data repositories.

Despite the sufficient number of IDA methods, the priority is gradually shifting towards logical algorithms for searching for cause-and-effect rules in data. They are used to solve the problems of forecasting, classification, pattern recognition, database segmentation, extracting «hidden» knowledge from data, data interpretation, establishing associations in databases, etc. The results of such algorithms are effective and easy to interpret.

In addition, the main problem of logical methods for detecting patterns is the problem of sorting through options in an acceptable time. Known methods either artificially limit such a search (KORA, WizWhy algorithms), or build decision trees (CART, CHAID, ID3, See5, Sipina algorithms, etc.), which have fundamental limitations in the efficiency of searching for causal rules. Other problems are related to the fact that known methods for searching for logical rules do not support the function of generalizing the above-mentioned rules and the function of searching for the optimal composition of such rules. A successful solution to these problems can be the basis for new IDA methods and related developments. The main advantage of using artificial neural networks is the ability to solve various non-formalized problems. The use of such networks for data analysis is advisable, since these networks are capable of approximating functions, learning and improving their own structure, provide a low probability of error under the conditions of correct initial setting of network parameters, as well as the ability to analyze even in the

presence of incomplete and noisy data. At the same time, it is relatively straightforward to model situations by feeding data into the network's input and evaluating the output it generates. The proposed algorithm for solving the task effectively captures the process of conducting the required analysis.

REFERENCES:

1. Balogh, E. (2015). Improving Diagnosis in Health Care. National Academies Press (US), Washington. https://doi.org/10.17226/21794

2. Beley, O. & Chaplyha, V. (2017). The application of neural networks for the intelligent analysis of multidimensional data, 2017 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkov, Ukraine, 400-404. https://doi.org/10.1109/INFOCOMMST.2017.8246426

3. Chao, Z., Wentao, L., Huiyan Z., & Tao, Z. (2024). Advances in Intelligent Data Analysis and Its Applications Special Issue Advances in Intelligent Data Analysis and Its Applications that was published in Electronics, January, 542. https://doi.org/10.3390/books978-3-03928-616-4

4. Glover, E. (2022). Artificial Intelligence https://builtin.com/artificial-intelligence (available: 28.08.2025)

5. Lyfar, V., Lyfar, O., & Zynchenko, V. (2024). METHODS OF INTELLIGENT DATA ANALYSIS USING NEURAL NETWORKS IN DIAGNOSIS, IAPGOŚ, 2/2024, 109-112. http://doi.org/10.35784/iapgos.5746

6. Narkevich, A., Vinogradov, K., Paraskevopulo, K., & Grjibovski, A. (2021). INTELLIGENT DATA ANALYSIS IN BIOMEDICAL RESEARCH: ARTIFICIAL NEURAL NETWORKS. Ekologiya Cheloveka (Human Ecology), 28(4), 55-64. doi:10.33396/1728-0869-2021-4-55-64

7. Vivchar, O., Tyukhtenko, N., Mykhailyshyn, L., & Korovchuk, Y. (2024) NETWORK MODELING OF ASSESSING THE IMPACT OF THREATS TO THE ECONOMIC SECURITY OF LOGISTICS OPERATORS. Economics, Ecology, Socium, Vol. 8, No.4, 89–98. DOI: https://doi.org/10.61954/2616-7107/2024.8.4-8 URL: https://ees-journal.com/index.php/journal/article/view/275

8. Why Recurrent Neural Networks (RNN) Dominate Sequential Data Analysis. (2024) March 29. https://shelf.io/blog/recurrent-neural-networks/ (available: 02.09.2025)

ПРАКТИЧНЕ ЗАСТОСУВАННЯ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ В КОНТЕКСТІ НЕЙРОННИХ МЕРЕЖ: БЕЗПЕКОЗНАВЧИЙ ПІДХІД В СИСТЕМІ ПРОПАГАНДИ

ВІВЧАР Оксана, СОБКО Ольга, МУЖИЛІВСЬКИЙ Назарій
Західноукраїнський національний університет

У статті досліджується практичне застосування інтелектуального аналізу даних у контексті нейронних мереж із позицій безпекознавства. Акцент зроблено на зростаючій актуальності використання штучних нейронних мереж як інструменту оброблення великих обсягів інформації, що виникають унаслідок автоматизації технічних і управлінських процесів. Визначено основні методи інтелектуального аналізу даних, проаналізовано їх переваги та обмеження, а також розкрито потенціал їхнього використання для забезпечення інформаційної та економічної безпеки. Обґрунтовано доцільність застосування нейронних мереж для вирішення складноформалізованих завдань, що потребують глибокої обробки, узагальнення та прогнозування даних. Розроблено підхід до побудови нейронної мережі для відбору даних з поетапним виокремленням процесів попередньої обробки, навчання та класифікації. Запропоновано алгоритм побудови найпростішої рекурентної нейронної мережі та наведено схему її архітектури. Визначено практичні напрями реалізації систем інтелектуального аналізу даних у сфері безпеки, зокрема для виявлення аномалій, оцінки ризиків і прогнозування загроз у корпоративних системах. Підкреслено міждисциплінарний характер інтелектуального аналізу даних, що поєднує статистичні, кібернетичні та математичні методи з інноваційними інформаційними технологіями. Особливу увагу приділено питанням оптимізації архітектури нейронних мереж, мінімізації помилок і підвищення швидкодії процесів навчання. Визначено перспективні напрями подальших досліджень, пов'язані з використанням гібридних моделей, формуванням теоретичних основ для інтелектуального аналізу даних та розробленням алгоритмів виявлення причинно-наслідкових зв'язків у великих масивах інформації. Зроблено висновок, що застосування нейронних мереж у сфері безпеки сприяє підвищенню ефективності систем аналізу, моніторингу й протидії сучасним загрозам.

Ключові слова: інтелектуальний аналіз даних, нейронні мережі, штучні нейронні мережі, рекурентна нейронна мережа, безпека підприємств, економіко-математична модель, пропаганда.